A SYSTEM APPROACH TO THE COUNTERMEASURES PROBLEM

W. P. Tanner, Jr.

September, 1960

IP-457
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF ILLUSTRATION</td>
<td>v</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 The Basic Problem</td>
<td>3</td>
</tr>
<tr>
<td>2. THE COMMUNICATIONS PROBLEM</td>
<td>14</td>
</tr>
<tr>
<td>2.1 Coding</td>
<td>23</td>
</tr>
<tr>
<td>2.2 Decoding</td>
<td>31</td>
</tr>
<tr>
<td>2.3 Use of Redundancy</td>
<td>37</td>
</tr>
<tr>
<td>2.4 Summary</td>
<td>42</td>
</tr>
<tr>
<td>3. THE RADAR PROBLEM</td>
<td>43</td>
</tr>
<tr>
<td>3.1 The Assumed Signal Evaluation</td>
<td>49</td>
</tr>
<tr>
<td>3.2 Receiver Efficiency</td>
<td>54</td>
</tr>
<tr>
<td>3.3 Operator Efficiency</td>
<td>54</td>
</tr>
<tr>
<td>3.4 The Capacity of a Signal to Lead to a Correct Choice among One of M Alternatives</td>
<td>59</td>
</tr>
<tr>
<td>3.5 Estimating M</td>
<td>62</td>
</tr>
<tr>
<td>3.6 Summary</td>
<td>64</td>
</tr>
<tr>
<td>4. THE COUNTERMEASURES PROBLEM</td>
<td>65</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>77</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>85</td>
</tr>
<tr>
<td>DISTRIBUTION LIST</td>
<td>87</td>
</tr>
</tbody>
</table>
# LIST OF ILLUSTRATION

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Basic Communication System</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>Rate of Information per Second as a Function of Signal Energy when Signal-Noise Power Ratio Is Constant</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>Rate of Information per Second as a Function of the Ensemble Size with Signal Energy Constant</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>Map Used as Visual Aid to Redundancy</td>
<td>41</td>
</tr>
<tr>
<td>5</td>
<td>Block Diagram</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>Probability of a Correct Choice as a Function of $2E/N_0$ with the Number of Alternatives as a Parameter</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>Block Diagram Illustrating Points of Measure in Radar Evaluation</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>$2E/N_0$ as a Function of Range when $2E/N_0 = 1.00$ at Range 1.00</td>
<td>52</td>
</tr>
<tr>
<td>9</td>
<td>Block Diagram Conceptualizing Experimental Design</td>
<td>56</td>
</tr>
<tr>
<td>10</td>
<td>Estimate of Human Observer's Performance Compared to the Ideal</td>
<td>61</td>
</tr>
<tr>
<td>11</td>
<td>$2E/N_0$ Required to Meet Criterion as a Function of Uncertainty</td>
<td>63</td>
</tr>
<tr>
<td>12</td>
<td>Block Diagram Illustrating Basic Countermeasures Problem</td>
<td>67</td>
</tr>
</tbody>
</table>
A SYSTEM APPROACH TO THE COUNTERMEASURES PROBLEM

1. INTRODUCTION

Anyone engaged in a countermeasures program will recognize at once the complexity of the problem before him. He is faced with equipment which he must evaluate in terms of both theoretical and experimental considerations. Conditions vary with the environment, and evaluative techniques which are satisfactory in one particular environment may be next-to-useless in another. Besides, there is the added complication of the human factor. The exact performance of a human being cannot be predicted with absolute certainty, and one must take into consideration the unreliability of the human component in any relatively flexible system. These problems are not easily solved. Indeed, in the past there has frequently been disagreement over the question of how the solutions should even be approached. Should we approach the matter from the point of view of theory, carefully stating the problem, controlling the environment, and collecting laboratory data for the evaluation of countermeasures systems? Or should we proceed from a practical point of view, considering as well the irrelevant variables found in field tests for the evaluation of equipment under specific conditions?

The present paper adopts the former approach for a variety of reasons. The distinction between theory and practice is sometimes
more apparent than real, for theoretical developments frequently have highly significant practical aspects. Such seems to be the case in the countermeasures problem. Evidence already gathered seems to point to the conclusion that only by approaching the matter from the theoretical point of view can the relevant relationships between the whole system and its parts be adequately studied. What we wish to do is to get to the fundamental nature of the problem, uncomplicated by irrelevant variables, so that we may both understand the functioning of a system and develop means for its evaluation. Such an approach requires that the subject be viewed in the light of recent communication theory.

Admittedly, such theory is extremely complicated and by no means easy to understand. It cannot necessarily be assumed to be in the general background of even very well-educated men. For this reason, the present paper assumes no knowledge of communication theory on the part of the reader. For those familiar with this area, the discussions which follow may seem overly simplified. The matter is handled in this way, however, so that any intelligent reader will be able to follow the discussion once he has become acquainted with the fundamentals of the theory. The paper attempts, therefore, to outline and explain the concepts from communication theory relevant to the present discussion in the hope that by this means the reader will come to see the significance of these new concepts for the evaluation of countermeasures programs.

It will be noted at once that the paper has a strong psychophysical bias. This orientation results from the primary field of interest of the author. The psychophysical viewpoint, however, is not the only, nor perhaps the most important, one presented in the subsequent
discussions. In the countermeasures problem we are faced with the task of evaluating both men and equipment. Fundamentally the task is the same whether we evaluate one or the other in terms of the whole. As we shall see below, it sometimes makes no difference whether a component under discussion is human or mechanical. It is the function of that component in relation to the system as a whole which is of primary significance. The reader, therefore, should keep in mind that we are dealing with systems. The psychophysical matters, though important in themselves, are also useful in illustrating a problem which has more general ramifications than psychophysics alone.

1.1 The Basic Problem

In any communication or detection system, two types of general problems are likely to appear. The first, and most easily recognized and understood, is the kind that relates to the particular components individually. Such problems may be isolated with greater or less ease and referred to particular specialists for solutions. Thus, the problems caused by the individual human component may be referred to the psychologist, those of the electronic components to the engineer. In many instances, however, a second kind of problem appears—that which concerns the relation of the individual part to the system as a whole. Thus, to use an example from psychophysics, the inclusion of a human component in a communication or detection system introduces not only the problem of human limitations, but also the problem of those limitations as they are related to factors outside the human being himself, most particularly, the lack of reliability of the human component in a flexible system. If the purpose of the equipment can be made specific enough, we are perhaps better off with automatic equipment, but in most
systems, which cannot be designed for a completely specific purpose, we must include a human being.

The exact nature or functional structure of this human component depends, in part, upon the way in which it is incorporated into the system. Human beings are extremely complex creatures, and unlike electronic equipment, cannot, at least at present, be represented by a schematic diagram composed of filters, amplifiers, differentiators, or the like. To be sure, there are specific tasks wherein the tie-in to the system and the purpose are so precisely specified that one fundamental schematic diagram of the human component may be used. If the task is changed, however, or if the tie-in to the system is changed, many of the parameters of the diagram most nearly representing the human component must also be changed. In some cases, it may even be necessary to change the basic functional structure represented in the diagram.

The underlying cause of this condition is that the human being is a self-evaluating and self-adjusting system—a fact which leads to both disadvantages and advantages. On the debit side, a number of major problems occur when an outside observer tries to predict the performance of a human being, particularly if the outside observer has difficulty in estimating the precise environment in which the system is operating. Suppose, for example, he requires that the human component handle more factors than he is capable of, or demands a storage capacity on the part of the operator beyond the limits of human memory. Situations like these would seriously affect the efficiency of the system or lead to its malfunction, simply because the human being could not operate according to specifications. Thus, the dangers incurred with the inclusion of a human component are easy to visualize.
Yet, although such problems are major ones and can seriously affect the performance of the system, these difficulties of prediction are not entirely disadvantageous. On the credit side, the incorporation of a human being into a system can lead to real advantages. An intelligent component can adjust if the original estimate of the environment is inaccurate. He can change with the situation by adjusting himself or the apparatus (or both) to fit the unforeseen conditions. Indeed, he can sometimes perform better than a mechanical component. Suppose, for example, that a mechanical device is set to receive a signal at precisely 1000 cycles per second. If the signal comes through at 1100 cps, the system fails completely. A human being, on the other hand, can react to the unexpected environment. If there is time, he may adjust to the new situation, manipulate his equipment, and receive the signal at the transmitted frequency. His behavior may, in effect, change the design of the system. He may thus perform better than specific automatic equipment because he adds a flexibility, a means of adjusting to the unexpected, beyond the capacity of most machines.

Because he can make up for system deficiencies, the introduction of an intelligent component can be of real importance to the engineer. The inclusion of a human being may permit certain design decisions to be delayed until the system actually goes into operation, for at that time the human operator may make the necessary adjustments to the specific conditions he finds. In certain cases, therefore, the design need not be made too specific, for in time the human component may adjust to the actual—and perhaps unpredicted—environment. He may even make the system work better than it would if specific design features had been incorporated. In effect, the system is deliberately
made less specific than it could have been so that the human component may make changes when more information is available upon which to base a design decision.

Perhaps an example will aid in clarifying this point. It is drawn from the theory of signal detectability, and a number of the concepts and terms may well appear new and unusual. Nonetheless, it illustrates the point well, and the explanation will introduce the reader to some concepts that will be useful throughout the paper. According to this theory, the cut-off point for accepting a signal (perhaps the bias on a thyratron) is optimally established on a knowledge of a number of parameters which define the particular environment in which the equipment is operating—parameters which can vary considerably from one situation to another. Stated flatly, these parameters are, among others, the \textit{a priori} probability of the signal occurring, the signal energy, and the values and costs associated with possible alternate decisions. Each of these concepts requires considerable explanation.

The way in which these variables play their roles can perhaps be demonstrated by the following illustration, which is not technical, but is quite graphic. Suppose that at the end of a busline frequently used late at night by nurses coming off the late shift at a hospital, there is a dark and deserted stretch of street lined with bushes and weeds. For many years girls have walked along this stretch completely unmolested. Should one of them hear a noise in the bushes, it would be accepted as a signal that a small animal is scampering away (the \textit{a priori} probability—the probability before the event—is high that that is all it is). The girl is not frightened, for it seems that little is to be gained by panic (value is low), and she may be laughed
at (the cost is high) for running at every noise. Suppose further, however, that a new factor is added to the situation. One night a girl is murdered. Once this fact occurs, the parameters change. It is highly likely that many reports will come in of girls running home in panic pursued by a molester, perhaps imaginary. What has happened? Because of the murder, the *a priori* probability, in the minds of the girls, of the signal meaning "murderer" becomes high. That is, to them any noise is likely to denote "murderer lurking near." This signal is heard more frequently (though the noise is perhaps of only very low intensity), and the values and costs have changed. If the signal is accepted, as it is likely to be (that is, "murderer is there"), the girl runs because the value of such action ("escape with my life") far offsets any possible cost. A change has occurred in the system because of the new information (the murder) which has been introduced.

To return to the engineering example, we may now see the problem the engineer faces in designing a system. The parameters vary from situation to situation, just as they did in the nurse example. Therefore, in designing a fixed bias into a system, the engineer must assume an environment, or set of environments, which leads him to elect that specific bias. But at the same time, obviously, there are likely to be many environments for which his choice is a very poor approximation. One new fact, as the murder of the nurse, fed into the system can change the parameters so markedly as to make the system totally unsuited for the environment in which it is asked to operate. And of course the engineer has no way of knowing what factors may crop up to bring about such a change. He may prefer, therefore, to provide for an adjustable bias in his system.
He would like to design a system which operates initially on the design engineer's best guess of the values of the relevant parameters in the environment. During its operation the system studies the environment, making new estimates of these values based on the information it collects. It continually modifies its basis of operation to conform to the most recent estimates of the environment. Since at the present time, however, it is a difficult problem to design these properties into automatic devices, the engineer must assign some of the functions of the system to a human being. With present knowledge he can in no other way achieve the flexibility and versatility required. He assumes, therefore, that the human component can introduce into the system the necessary self-evaluation and self-adjustment in a satisfactory manner. That assumption, however, is based on little evidence and much faith. Without question, the human component performs these functions. That he performs them satisfactorily is by no means so certain.

Until quite recently this whole question of human performance had not been deemed sufficiently important to justify major consideration, mainly because it had not been necessary to demand the best possible use of human components in both communication and radar systems. For example, the relatively small amount of traffic in communication systems allowed plenty of bandwidth to be assigned to individual networks, and in radar systems there has ordinarily been sufficient time for human beings to react to signals. Under such conditions, it was not at all necessary to demand optimum use of the human component.

Recent developments, however, indicate that in the not-very-distant future systems which employ human beings will have to make much greater use of their capabilities, mainly because of reduced time and
increased traffic in radar and communication systems. Consider, for example, the radar problem. The speed of modern missiles and counter-missiles is such that recognition and interception must occur in intervals of time beyond the fastest of human reactions. To use a commonly recognized illustration, the problem of the human operator is similar to that of two pilots in jet planes approaching each other while flying on visual observation alone. Speed has become so great, and reaction time requirements so short, that the planes may collide before the pilots are even aware of any danger. To be sure, these are extreme cases in which the unaided human being can do nothing. In other cases where more time is available, the human capabilities will have to be taken into much more careful consideration.

The problem exists as well in communication systems. Traffic has increased to such an extent that extravagant assignment of bandwidth to individual nets is no longer possible. The nets must operate on narrower bands, a condition which entails some serious problems, not all of which will be immediately apparent to the reader, for they involve some concepts from information theory which will be handled in more detail in Chapter 2. Nevertheless, the general nature of the problem may be mentioned here; more detailed analysis of the concepts will follow. It is hoped that as we proceed, the matter will become increasingly clear.

The first and most obvious effect of the narrower bands is the reduction of the information-carrying capacity of the nets, an effect which entails the need for more efficient use of the assigned channels if the same rate of information flow is to be maintained. A number of consequences come to mind. Suppose we assume a human being
speaking through a channel. The code he uses may have to be more efficiently constructed in terms of the capacity of the channel.

The word "code" is not used here in its most usual meaning. Any accepted means for transmitting information—language, flashes of light, dots and dashes, even tone of voice—may be classed as a code. If the human voice employing language is transmitted over a narrow channel, we may have to specify the tone of voice or the limits of high and low inflections (and this is part of the code) which must be used if the message is to be received with minimum uncertainty. Thus, the codes must be carefully designed, for we cannot afford to depend upon communication systems for which our expectations are unrealistic.

More efficient utilization of the channel capacity will, of course, increase the susceptibility of the system to jamming. Jamming reduces the capacity of the channel. Since the bands are smaller, less noise power will be needed to jam the message (though naturally it will be harder to introduce). To counter this susceptibility, it will be necessary to make an intelligent use of redundancy. That is, we will have to be careful that those elements which do not add information do, in effect, serve as error-correcting elements, thus increasing the certainty with which the message is received. Notice, for example, the redundant letters in an English sentence; we can recognize the words even if letters are omitted. A wide band gives us a kind of wasteful redundancy—much more than we need. Since we will be forced to avoid this type, we will have to use that which remains to reduce the danger of equivocation—of uncertainty on the part of the receiver—which increases with the narrower bands. We may have to specify, for example, detailed standard operating procedures to fit existing conditions so
that we may be sure that redundancy is built into the system.

In the future, therefore, it appears that more attention will have to be directed toward the problem of using the capabilities of the human component more efficiently in those situations in which they are to be used. The knowledge basic to this consideration will come from psychophysical experiments, probably performed within the framework of models of such recent development that their potential has not yet been generally realized—statistical decision theory and the theory of signal detectability. This is an area of research which at present has only a brief history. It dates primarily from Claude Shannon's paper on the mathematical theory of communication, published in 1947. This and other studies, as will be seen in subsequent sections, have furnished the scheme of the basic communication system and certain fundamental concepts pertinent to the present problem. Although at first glance the concepts seem difficult (the vocabulary, in particular, is not easy to master), they are of such importance that anyone working in a countermeasures program should be—and can be—aware of what they imply. For despite its short history, the theory has already demonstrated a potential to furnish some of the required knowledge.

We shall begin, therefore, with a brief outline of the pertinent concepts in information theory, and analyze the communication, radar, and countermeasures problems in terms of these concepts. All three problems are included for a definite reason. Before the countermeasures problem can be fully understood the reader must be thoroughly familiar with the natures of the target systems themselves. Since the communication and radar analyses illustrate problems inherent in the target systems, each is treated at length before the analysis of the
countermeasures system. The communication problem is basically a recognition one. It involves two people in a communication game in which the message ensemble can be known in advance. The radar problem differs in that it also involves detection, and the message ensemble (the targets) cannot be as precisely known in advance. Only when these problems are clarified can the countermeasures problem be properly analyzed.

Analysis of these problems will reveal a number of areas requiring research. Some of the problems are quite specific. Thus, in the communication problem, as we shall see below, this question arises: information already gathered indicates that from the decoding standpoint, sequential observation (observation over a flexible time interval until the observer achieves a level of confidence that he has received the message correctly) is more efficient than fixed time observation. Since this is true, it becomes important to know the extent to which a human operator can act as a sequential observer. Other areas for study, however, include those in which even basic knowledge is lacking, for example, the manner in which a human operator on the receiving end of a communication channel stores and decodes messages (that is, reconstructs the transmitted message from the signal he has received).

Each of the three general problems (i.e., communications, radar, and countermeasures) is stated within the theoretical framework of communication theory and the theory of signal detectability. By establishing a single consistent framework it is possible to treat complex problems involving more than one of these systems. The
framework leads to the establishment of those relevant measures upon which the evaluation of these complex systems and their components can be based. It is, further, the framework within which basic studies should be conducted for the purposes of increasing the scientific knowledge leading to future advances in the countermeasures program.
2. THE COMMUNICATIONS PROBLEM

Before any effective countermeasures can be taken, the fundamental nature of the communications system must be clearly understood and the problems relevant to the general area of communications carefully analyzed. The problems we are concerned with are inherent in the systems themselves, a fact which must be constantly kept in mind throughout the discussion. Only when we understand the system can we expect to employ effective countermeasures. The analysis of the problems presented here, therefore, will attempt to isolate areas where further research is needed for a fuller understanding of the general system. In this section of the paper special problems of coding, decoding, and the use of redundancy will be analyzed with a view toward suggesting a number of areas in which further research and study are especially desirable. Since these terms, however, are used in a very special sense, a considerable amount of background will be necessary.

An understanding of the problems will require some knowledge of information theory, for it is in terms of such theory that the problems will be presented. The basic communications system and its important components are well illustrated in the block diagram upon which Shannon based his fundamental theorems, and the theorems themselves (although they are statements of averages which cannot be indiscriminately applied to realizable systems) are useful in helping us formulate the problems. A word of warning should perhaps be added here. The application of the theory depends upon large samples which permit statistics to apply. It is a theory of averages, and, although we can make predictions about a system on the average, we cannot predict the outcome of any individual event.
An illustration may help. In the case of the decay of radio-active material, we can predict that within a certain length of time, a portion of the electrons will escape from the pile. We cannot predict, however, what an individual electron will do. In a similar fashion, in information theory we may establish theorems based upon a large number of messages, but the theorems may not furnish the basis of stating the fate of any one in particular. The problem causes difficulty in information theory only because we are concerned with individual messages, whereas we are not concerned with individual electrons. If the reader will keep this warning in mind, however, he should have no trouble with the theory.

Shannon's diagram, modified only by the addition of a feedback channel, is reproduced in Fig. 1 and explained below. The system presented here is an abstraction - a general system not descriptive in detail of any specific system. Nevertheless, its very generality is the basis for its usefulness, for it describes - abstractly - any communications system, from high speed electronic devices to two friends engaged in conversation, from a radio or television broadcast to a pair of Indians sending smoke signals. In each case, the basic components are the same.

There is an information source which selects a desired message from a set of possible ones, each with a probability that it will be transmitted when transmission occurs. The transmitter encodes it into a suitable code. That is, it changes the message into the signal which it sends over the communications channel, any medium - a pair of wires, a band of radio frequencies, a beam of light - that can be used to transmit the signal to the receiver. The receiver accepts the received signal - the transmitted signal plus noise, any addition or change (distortion or error) not intended by the transmitter - and decodes it. That is, it
FIG. 1 BASIC COMMUNICATIONS SYSTEM

- Information Source
- Transmitter
- Receiver
- Noise Source
- Feedback Channel
- Message
- Signal
- Received Signal
- Destination

16
reconstructs the message from the signal and passes it on to the destination, the person or thing for whom the message is intended.

Perhaps an example will illustrate the fundamental nature of the abstract system. Consider two friends engaged in a conversation. The speaker's brain is the information source; his vocal system, the transmitter. His words (the message) are transformed into varying sound pressures (the signal) and transmitted over the air (the channel). His friend's ear and auditory nerves are the receiver; his brain, the destination. Similar examples could be constructed from any situation in which communication is taking place. The general system describes them all equally well.

From the point of view of this section of the paper, the most significant parts of the process are the coding and decoding of the message; that is, we will be most concerned with the problems that arise when the message is coded (changed into the signal) and decoded (reconstructed from the signal-plus-noise that has been received). These problems are the same whether the coding involves language, binary digits, flashes of light, or current passing through a cable. Let us turn, then, to the application of Shannon's theory to these problems.

Around the block diagram Shannon constructs a coding theory, statistical in nature, which will be useful in our discussion of coding problems. To be sure, his theory must be used with care, for, as Shannon himself has noted, the hard core of the theory is a branch of mathematics, a strictly deductive system. The theorems he presents are statements of statistical averages expected of communications systems exhibiting stable statistical properties. By such a system we mean that we are dealing with infinite series of events, all of which have the same experience
entering into them. The system is not altered by experience, nor does it change as it goes along, unless such change is built into the system by some general statistical rule. Let us also stress again that we are dealing with averages, which, like all averages, cannot be applied indiscriminately to particular situations. Shannon's fundamental theorem, for example, assumes transmission forever. Such a principle, based upon infinite time, cannot be applied to realizable, finite systems unless one is well aware of the dangers lurking in statistical statements.

Still another difficulty has to be mentioned. The vocabulary of information theory is somewhat hard to master, not because the words are difficult, but because they have common as well as specialized meanings. We have already noted the rather special sense in which the word code is used here to mean something much less specific than the ordinary acceptation of the term. Other words are even more difficult to grasp - information, for example, which is not meaning, but rather a measure of one's freedom of choice when he selects a message. Confusion can result in one's mind because a word like information has these two meanings - one in general use, one in terms of information theory - both of which are needed in discussing communications. The reader must constantly keep in mind, therefore, how these words are being used.

Nevertheless, despite these limitations, Shannon's theorems are useful in helping to formulate the problems, and will therefore be used throughout this section, but especially in the analysis of the problems of coding. No attempt will be made to prove these theorems. The proofs require a considerable amount of mathematics and are clearly beyond the
scope of this paper. The interested reader may consult Shannon's work itself. All we can hope to do for the purpose of this paper is to state and explain the theorems so that the reader may understand the concepts involved. Most pertinent to this discussion are the four concepts detailed below.

(1) The information flow in a communication system is the same whether viewed from the standpoint of the transmitter or the receiver. Let us repeat once again that the reader must keep in mind, both here and throughout the entire discussion, that information, as used here, is not meaning, but rather a measure of one's freedom of choice when one selects a message. Perhaps we can explain the term best by these means. In information theory, the word entropy is used to express the degree of randomness in a situation. It is measured logarithmically and is expressed in terms of the probabilities involved. The more equal the probabilities, the greater the entropy. Information, then, is a measure of the reduction of entropy, or of uncertainty, in a communications system. What the theorem is saying in effect is that the channel can reduce uncertainty by so much, and it does not matter from which end you look at it.

An example or two may help to illustrate this concept. Let us assume a channel transmitting the digits 0 and 1, with each digit equally likely to be transmitted. From the transmitter's point of view, the rate of information flow is measured by the change in his ability to predict the received digit before he knows the symbol to be transmitted and after he knows it. From the receiver's end, the rate of information flow is the change in his ability to predict the symbol which is transmitted before and after the signal is received. In our assumed channel, the 0's and 1's may be mapped into two things each. Let us say that a 0 may be
A or B, a 1 either C or D. 0 and 1 are equally likely to be transmitted, and the received digits are A, B, C, and D.

From the transmitter's end of the system, before he knows the signal to be transmitted, he knows that the receiver will receive one of four digits, each with a probability of .25. After he knows the signal to be transmitted, he knows the receiver will receive one of two digits, each with a probability of .5. His uncertainty has changed from two bits to one bit; it has been reduced by one bit. From the receiver's viewpoint, before he knows the signal, he knows either a 0 or a 1 will be transmitted, each with a .5 probability. He has, therefore, one bit of uncertainty; after the message has been received, the uncertainty has been reduced by one bit. Now notice what has happened. The transmitter's uncertainty has been reduced from two bits to one bit, the receiver's from one bit to none. In each case uncertainty has been reduced by one bit. Since the information flow is measured by the reduction of uncertainty, it is clearly the same no matter from which end of the system we look into it.

Let us examine another example. Consider a system transmitting the binary digits 1 and 0. For the sake of a simple illustration, we will assume that it is a symmetrical system. The probability before selection of the digits is .5 that either will be selected. This probability is the same from the point of view of either transmitter or receiver. Suppose further that once a 0 is transmitted, the probability is .9 from the transmitter's point of view, that a 0 will be received. After the perturbed signal (signal plus noise) is received, the probability is .9 from the receiver's point of view, that the 0 is correct. For ease of calculation, we have assumed a symmetrical system. The same holds true for an asymmetrical system. The calculations, however, are so involved
as to be out of place in a paper of this type. Actually, the point should be clear without them: the channel can reduce uncertainty by so much, and it matters not at all whether we view it from the transmitter's end or the receiver's end.

(2) There exists a code which leads to arbitrarily small error and still permits information flow at a rate nearly that of the capacity of the channel, while any code which attempts to transmit at a greater rate must lead to an equivocation¹ at least as great as the difference between that rate and the capacity of the channel. This means that so long as the capacity of the channel is equal to or larger than the entropy of the source of the messages, we can transmit with as small a specification of error as we choose, so long as the transmission process continues over an infinitely long period of time. With time unlimited, we can build into the system error-checking devices. The longer the time allowed for this purpose, the more effective these checking devices can be.

Consider some of the means by which this can be accomplished. One method of error correction is simple repetition. The message, for example, may be repeated once - 1389, 1389 - and we have a redundancy of 50% in the total message. Or we might repeat the message three times - 2469, 2469, 2469 - and the redundancy is 67%. We might prefer to add the sums of the digits - 2469 21 (redundancy is 33%), or perhaps only the last digit of the sum - 2469 1 (redundancy is 20%) - to show that the message is correct. The principle is one familiar to those who have had experience with legal language. In an act of sale of a piece of property

¹Equivocation refers to the average uncertainty existing after the transmission.
a lawyer may write that the seller agrees to "grant, bargain, sell, convey, assign, transfer, and set over" that property to the buyer. What he is doing in effect is building redundancy, an error-checking device, into this message to make sure that it is not misunderstood. We can easily see, therefore, that with infinitely long time for transmission, we can send messages with arbitrarily small error, so long as the capacity of the channel is equal to or larger than the entropy of the source of the messages.

The above examples are merely intended to illustrate the role of redundancy. In the first example, that of repetition, no new information is transmitted after the first transmission. The second example again has this feature; the transmission of the sum is pure redundancy and serves only to detect errors. In the third, that of the legal language, each word carries some of the information included in every other word and some that is not. A single word could be coined to cover this set of words. None of these examples pretends to illustrate a use of redundancy which reduces error without reducing the rate of information flow.

If, on the other hand, we attempt to transmit at a rate greater than the capacity of the channel (the capacity in this case is smaller than the entropy of the source), we incur an equivocation equal to or greater than the difference between the capacity and the rate. What we are mainly concerned with here is the relation between the capacity of the channel and the rate of transmission. Thus, if we vary the channel capacity (as we do when employing countermeasures) so that the channel capacity is less than the rate of transmission, the theorem tells us that there must be equivocation, or uncertainty, when the message is received.
The remaining points may be very briefly made:

(3) The rate of information flow from transmitter to receiver in the system illustrated in Fig. 1 is restricted only by the capacity of the forward channel.

(4) The feedback channel can be employed to reduce the error rate of the information flow from transmitter to receiver, but not to increase the capacity of the forward channel.

Other concepts, especially from statistical decision theory, will also be employed, but since they are related primarily to problems in decoding, we will postpone discussion of them to that section of the chapter and proceed at once to discussion of the coding problems.

2.1 Coding

From the discussion above it is obvious that the employment of redundancy is an important coding problem. There is a great deal of highly sophisticated mathematical effort being extended in this direction. The fact that this is not treated in detail in this paper is not intended to minimize the importance of this work. It is extremely important, but requires more technical discussion than is possible in this paper. However, we are devoting the discussion to voice-communication channels, or subject matter not sufficiently understood to justify sophisticated mathematical treatment.

A number of coding problems deserving further study have been suggested by research conducted at The University of Michigan. Two in particular involve the use of a human being on the transmitting side of a communications system. These are interrelated, for they involve not only the question of coding procedures in voice communication so that we
may achieve the greatest information rate per unit time, but also raise
the question of how these procedures should be determined for a given
situation. We have already seen that a human being can add flexibility
to a communications system simply because he is a self-evaluating and
self-adjusting component. To apply this concept to the coding problem,
we may note that he has the ability to change the code through the use
of varied reading rates, changes in volume, shifts in inflections, and
the like. The problem, therefore, is how to put this flexibility to use.
In other words, what can the human operator do to adjust his coding pro-
cedures to achieve the highest information rate per unit time? And, more
important, can we devise a general rule which will tell the human operator
which of the changes to use in a given situation?

From this point of view it is important first to know the
maximum information rate achievable in any given system. Figure 2 illu-
strates this problem. It is derived from calculations which are beyond
the scope of this paper. The reader is asked, therefore, to accept the
graph as a true picture of the situation. The graph shows that, for a
fixed signal-to-noise power ratio in a digital channel, the information
rate per unit time varies as a function of the energy per symbol, or
equivalently, as the duration of the symbol. It applies to a channel
in which each symbol in a sequence is independent of every other symbol
in the sequence. The two lines illustrate the relation in situations
based upon two different assumptions. The first assumes that the signal
is known exactly; that is, if the signal exists, its waveform is known
exactly even to its position in time. The curve shows that the reduction
of error in such a channel by increasing the energy of the symbol can be
accomplished, in this situation, only at a cost of a loss of information
rate.
The other assumes that the signal is known only statistically; that is, that we know the waveform, but perhaps not its position on a fixed time scale, or perchance, the amplitude. In such a situation, the curve achieves a maximum point in time, a fact which suggests the need for further study in coding procedures. Is it possible, in a voice-communication channel, to achieve something like this maximum? Is it possible to plan the redundancy in a way which permits the reduction of error without paying an excessive price in reduction of information flow? As we have noted, by Shannon's fundamental theorem, there exists a code which can reduce error to an arbitrarily small amount without the loss of information rate illustrated in the first curve. By this theorem also, there exists a code which can lead to an approximation of the information rate indicated by the maximum of the second curve in Fig. 2, with arbitrarily small error. While it would be extremely optimistic to attempt to achieve these codes in voice-communication channels, their existence indicates directions in which progress might be sought.

This optimum code, however, as we have seen, requires infinitely long delays to achieve the arbitrarily small error rate, for the theorem assumes transmission over an infinite period of time. In situations in which a fixed amount of information is to be transmitted, or in which the messages are independent, the theorem may not apply. In such situations, correction must come within the message itself or not at all, for we frequently cannot afford the time needed to assure minimum error through correction devices built into the code. If the message contains information urgently needed at the receiving end, we cannot waste time purely for the sake of an arbitrarily specified reduction of error. Nevertheless, there may well be optimum codes achieving a minimum error rate in accor-
FIG. 2 RATE OF INFORMATION PER SECOND AS A FUNCTION OF SIGNAL ENERGY WHEN SIGNAL-NOISE POWER RATIO IS CONSTANT
dance with the restrictions placed on the communications task. In other words, of a number of possible codes, we would like to find the one which, though falling short of the ideal, still has a minimum error rate in terms of the finite task that the communications system is required to perform.

From this point of view, it becomes highly important to conduct studies, either theoretical or experimental, to establish maxima, or approximations of maxima, for curves such as that in Fig. 2 for situations in which the signal is known only statistically. When a human component is built into the system such studies would include research in voice communication. We would like to know, for example, the effect of reading rate in the establishment of such a maximum point. We should consider as well, other variations, such as change in volume, tone, and pitch in the human voice.

An example may clarify the problem. Suppose that a babble of voices is used as an interference signal in a communications system. When this sound is evaluated using an articulation test, that is, one with a fixed reading rate, it is found to be quite effective in interfering with the system. But when the person transmitting is permitted the freedom to vary his rate of speech, he can sometimes find a reading rate which will permit him to get through the babble and render the interference ineffective. Almost everyone is familiar with this problem from his experience with crowds at cocktail parties and in night clubs. Despite the hubbub, one can frequently find a code (a speed of delivery, a volume, or a tone of voice) which will permit him to converse with his companion. Despite this ability we cannot, of course, remove equivocation completely in finite time, but we can certainly optimize to get the best performance per unit time while accepting some error.
The second problem in coding procedure - the need for a general rule by which we can determine the particular code to use in a given situation - is illustrated by the data plotted in Fig. 3. Once again the reader is asked to accept the figure and the calculations that have gone into it without further development. In this figure, the information rate per symbol or per message is plotted as a function of message ensemble size wherein the signal-to-noise energy ratio per symbol is fixed. In this particular illustration, ensemble size is roughly analogous to vocabulary size in an articulation test, and each curve represents a different value of $2E/N_0$, a ratio which is so important to the understanding of both this and the radar problems that some explanation of it must be introduced here.

Stated briefly, the ratio $2E/N_0$ (where $E$ is the signal energy and $N_0$ is the noise power per unit bandwidth) describes, for the case in which the signal is known exactly, the separation between the means of two statistical distributions divided by the standard deviation, when these two distributions are along a unidimensional decision axis. One of these distributions is for the probability density of a measure on the decision axis when noise alone exists, while the second is conditional upon the existence of signal plus noise. We know further that for small values of $2E/N_0$ we get a large amount of equivocation with large ensembles, and a relatively small amount if the ensemble size is small. As we increase the value of $2E/N_0$, however, we decrease the equivocation in bits in large ensembles much faster than in small ones. (See Fig. 3 for an illustration of this relation.) For each value of $2E/N_0$, therefore, we have an information rate per symbol or per message for each ensemble size, and for each such value, there is an optimum
FIG. 3 RATE OF INFORMATION PER SECOND AS A FUNCTION OF THE ENSEMBLE SIZE WITH SIGNAL ENERGY CONSTANT
ensemble size for the rate of information flow.

The curves in Fig. 3 illustrate this fact. As mentioned above, each plots a different value of $2E/N_0$. According to the curves, for each signal-noise energy ratio, there is an ensemble size which leads to a maximum rate of information flow. Ensembles smaller than this do not have sufficient entropy to justify the capacity of the channel. In other words, the capacity of the channel is so great in relation to the entropy of the ensemble that a large amount of channel capacity is wasted. Conversely, ensembles greater than this are too large for the capacity of the channel. And as Shannon has shown (Ref. 1), if we attempt to transmit at a rate greater than the capacity of the channel, we incur too great an increase in error.

One is tempted, of course, to think further in terms of Shannon's fundamental theorem, and to search for the ideal code, one which will permit information flow at a rate near the channel capacity with arbitrarily small error. The theorem, however, assumes infinite time for transmission and cannot be applied to a finite system. It is based, moreover, on an existence proof. Although we know that the code exists, it is not defined. Yet despite these limitations, the concept remains useful, for it does state a bound, or standard, against which to compare performance. It represents an ideal in terms of which the efficiency of the real can be judged. It establishes as well a limit of expectations for the results of efforts to improve coding systems. In other words, it presents an ideal which we can perhaps approach, even though we can never actually reach it.

In voice communication studies, for example, we may find that through intelligent experimentation we may be able to determine general
rules useful in establishing the best of a set of procedures. Consider, as an illustration, the reading rate established in a Standard Operating Procedure. We would like to be able to determine the optimum reading rate based upon the actual situation encountered. In other words, since the human component adds flexibility to the communications system in that it allows the system to adapt to unforeseen circumstances, it would be valuable to establish rules upon which individual standard operating procedures could be based. Any given procedure, therefore, would not be standard for all situations, but for particular kinds of situations. Such knowledge should help increase information flow and decrease equivocation. It should enable us to approach the best possible conditions.

2.2 Decoding

A second general area in which further study is essential is that of decoding and, most especially, the function of the human being on the receiving end of a communications system. To understand the problems raised here, one must be somewhat familiar with not only the concepts of Shannon's theory, discussed above, but also with three concepts from statistical decision theory. Let us first state them flatly and then proceed to an explanation of each. They are: (1) Woodward and Davies' (Ref. 2) assertion that a statement of a posteriori probabilities contains all the information in the received signals; (2) Van Meter's (Ref. 3) recently advanced conclusion that decisions on parts of a message should be made to preserve as much information as possible, while decisions on the total message should be made in relation to some other criterion, such as the necessity for action; and (3) Wald's concept (Ref. 4) of sequential observation as more efficient
than fixed time observation in that the same error restrictions can be satisfied with less average time. In the following discussion, each of these ideas will be used as appropriate to the analysis of the problem.

Basic to the discussion is Woodward and Davies' concept. They have shown that a statement of a posteriori probabilities contains all the information contained in a receiver input. Remember, once again, that, as mentioned above, we are concerned with the question of probabilities. This concept merely restates the one made earlier - that it does not matter which end of the communication system we look through; the information remains the same. Once we decode, however, we incur the danger of losing information. Doubtful signals must be decoded with one meaning or another. As long as we have only the signal, there is a possibility of correction of error. Once the decision as to its meaning is made, however, information can and will be lost.

This concept is important because it gives us a general rule for the question of when to decode. According to this concept, one should decode only when a decision is necessary for some consideration beyond those involving information. It may be necessary, for example, for the information to be relayed to another station by a different means. The relay channel may not have the capacity to transmit the received signal. In such a case, the information must be decoded and sent. Or the nature of the message may be such - the presence of unidentified aircraft approaching in force - that we willingly incur the loss of information so that necessary action may be taken. In the latter consideration, the need for action quite obviously overrides the need to preserve information. In other cases, however, where no
such decisions are necessary, we should be mainly concerned with preserving information.

This conclusion is certainly consistent with the concept recently advanced by Van Meter. He presents the view that all efforts should be directed toward preserving information until the need for action arises. It may be necessary, for example, to make a decoding decision because of a storage limitation. If this necessity should arise, the decision should be based upon criteria designed to preserve as much information as possible. To preserve information should be the major concern in every situation except that in which action becomes necessary. Only then should the criteria be changed to base the decision on getting the best possible results for the action. In other words, until there are uses for the information, all efforts should be directed toward storage of information as such. When the uses exist, the efforts should be directed toward optimizing the use.

The importance of these concepts is apparent when we consider a human component on the receiving end of the system. There are many things about him that we need to know, most especially, how he stores and decodes the message. Does he decode unit by unit, symbol by symbol? In other words, is he like a novice trying to hold a conversation in a foreign language, translating each word as it comes up? Or does he, like a fluent speaker, preserve sentence after sentence in his mind, until he must act, that is, reply to the one to whom he is speaking? Does he, therefore, store information as far as possible, making decisions only when there is some compelling reason to do so? It is important to know how he acts, for what he does can seriously affect the efficiency of the system.
Suppose, further, that the time has come to act, that it is
time, for example, for him to transmit to another station that infor-
mation which he has so far received. Upon what criterion has he acted?
Did he base his decision, as suggested by Van Meter, with a view toward
preserving as much information as possible? If, on the other hand, he
has decoded so that action may be taken, has the decision been based
upon principles that will bring about the best possible outcome of the
action? These questions clearly indicate that there are many problems
raised by the inclusion of a human component in a communications system,
problems concerning which further study is essential.

Still another problem deserving further research is suggested
by Wald's concept that a sequential observation is more efficient than
a fixed-time observation in that the same error restrictions can be
satisfied in an average of half the time. This concept is a most impor-
tant one and requires considerable explanation. The following illus-
tration may help. Suppose a man has been invited to deliver a lecture.
If he is important enough, he might conceivably elect to put his lec-
ture on tape and have it played for his audience. To do this, however,
he will have to estimate in advance the level of knowledge of his audi-
cence, the conditions in the auditorium, their attentiveness, and many
other factors. He will have to make his recording on the basis of his
estimate of what the average conditions will probably be. If he speaks
in person, however, he can base some of his decisions on information he
collects as he speaks. He may be able to judge from the facial expres-
sions, nods of the head, etc., when each successive idea has been under-
stood. Thus, he dwells on each idea until he receives confirmation
from the audience that it is satisfied. On some of the ideas he may
spend a longer time than he would have had he taped his speech in advance. In others he may find he does not need to spend as much time. On the average, assuming that it is a long speech containing many ideas, the speech would require less time per idea if he delivered the talk in person than if he taped it in order to achieve the same level of understanding on the part of the audience.

If effect, the tape-recorder example is analogous to the problem involving fixed-time observation. With a pre-established code - one constructed in advance - we must work upon the principle of averages. We accept a time unit which, on the average, should give us a reasonable probability of certainty, though we recognize and accept the fact that error is present. It is like planning a stock-buying campaign in advance. We know that we will make errors somewhere in our estimates and calculations, so we must make our plans on the average if we are to have any success. We have to accept error. In a similar fashion, with fixed-time observation in a finite system, we must also accept a degree of error.

On the other hand, the sequential observation - and our speaker permits his audience to act sequentially - enables us to operate more efficiently. Instead of operating on averages alone, averages which may not be at all descriptive of a given situation, the observer makes use of information immediately available. Thus, instead of observing for a fixed length of time, he observes only up to the point at which he achieves a certain level of confidence. He continues to observe until that point is reached. To follow up our stock market analogy, with sequential observation, we would not plan the campaign in advance,
but would make our decisions on purchases and sales on the basis of day
to day information. We would have a clearer picture of the situation
in which we are operating and thus should be able to reduce error.

Perhaps the most significant fact about sequential observation
is that the observer has control of the length of observation time.
Once he has the information, or has achieved his degree of confidence,
he doesn't have to waste time by building in redundancy or error-
correcting devices. He can accept the signal and go on to receive
additional information. On the average, the observer who operates
according to Wald's concept cuts the length of time to the point of
removal of uncertainty to about half. Thus, he raises the degree of
certainty per unit time.

Since sequential observation is so significant, we may well
ask how this kind of process may be built into a communications system.
The block diagram in Fig. 1 contains the means which makes this kind of
observation possible - the feedback channel which serves the same func-
tion as the audience's reaction did in our speaker example. Only a low-
capacity channel is needed, since its sole function is to transmit to
the source the information that the receiver has now completed an obser-
vation. The receiver listens to the signal until he is in a position
to accept one of the symbols with a satisfactory degree of confidence.
At this point, he transmits over the feedback channel a single binary
digit which states that he is now ready to start the observation of the
next symbol. The capacity required of the feedback channel used in
this way is at most one bit per symbol. By using this means, the human
observer in a communications system can observe sequentially. And in
acting in this manner, he can satisfy the same error restrictions as in
fixed-time observation in an average of half the time.

Since sequential observation might increase the efficiency of a communications system to such a marked degree, and since, also, the human component may serve as such an observer, it is clearly important to know the extent to which the human operator can act as a sequential observer. It will be noted also that this type of observation adds a desirable kind of flexibility to the communications system since it allows design decisions to be delayed until the precise environment is known in which the system is operating. Studies in this area, therefore, should provide information for improving the efficiency of communications systems.

2.3 Use of Redundancy

A third problem requiring further analysis and research is that of the use of redundancy in a communications system. The term redundancy as used here is that "fraction of the structure of the message determined not by the free choice of the sender, but rather by the accepted statistical rules governing the use of the symbols in question." Varying degrees of redundancy may be built into any system. We have already seen in the general Introduction the amount of redundancy (actually about 50%) in the use of letters in the English language, and we have discussed above (in terms of the second concept from Shannon's theory) how redundancy can be built directly into a message. Such redundancy, of course, decreases the information transmitted, but adds measurably to the degree of certainty with which the message is received. It serves, in effect, as an error-correcting device.

Fundamental to this discussion of redundancy is the first concept from Shannon's theory quoted above: that the rate of information
flow is the same viewed from the standpoint of either the transmitter or receiver. This concept is very important in understanding the effectiveness of the performance of the total system. The reader should refer back to the first example given in the discussion of this concept, that in which the uncertainty of the transmitter was reduced from two bits to one bit, that of the receiver from one bit to none, so that in each case the uncertainty was reduced by one bit. In this case, the a priori uncertainty of the receiver was not matched to that of the transmitter. With the mismatch, there must be some equivocation at one end of the channel. It is only when the two ensembles are matched that equivocation can go to zero at both ends of the channel. Thus, the uncertainty of the receiver should be matched to the uncertainty of the transmitter, a matching that can be achieved only through the use of a priori information - that is, when the receiver has knowledge of the message ensemble.

This fact suggests the necessity of matching the ensemble of the transmitter with the ensemble of the receiver. Consider the following illustrations. Suppose the ensembles are completely unmatched, that an audience which understands only English suddenly finds itself being lectured to in Japanese. Since there is almost no relation between the languages, the message ensemble of the speaker is totally unknown to the listeners, and little or no communication takes place. The problem exists, moreover, even when the ensembles are partially matched. Suppose our lecturer speaks English, but has in mind a frame of reference for his words different from that of his audience (he is of a different religion or political party, or has a different social or economic view). It is possible that they may understand the words, but
may be puzzled or confused as to his meaning. He is using an ensemble that they do not understand, and his message, perhaps, does not get through to them at all. Everyone is familiar with such common misunderstandings, even in the conversation of friends of long standing.

Essentially, therefore, the problem is one of the use of a priori information, of giving the receiver knowledge of the message ensemble. In our latter example, the speaker may evaluate his audience in advance and adjust his ensemble to the average to be expected. Or in seeing the puzzled looks on the listener's faces, he may restate his ideas in other forms - in other words, he may change the ensemble to fit theirs and so get through to them. In still other instances, the audience itself may have acquired information concerning the point of view or frame of reference of the speaker. (Students frequently "case" their professors in this way.) Thus, they know in effect the message ensemble, and misunderstanding is less likely to occur. In a communications system, however, if the message ensemble is of considerable size, the memory of the human operator may be severely taxed.

It is important, therefore, to find means by which the human memory may be helped. To be sure, Miller (Ref. 5) has reported some interesting studies suggesting that human memory is quite restricted. Nevertheless, even though this may be true, there are means available which can certainly help. The human being does not have to rely upon unaided memory, for the use of memory aids can help him to a rather substantial and quite accurate memory. One such memory is the map used in communication studies at The University of Michigan.

This map is useful because it gives the receiver a priori knowledge of the message ensemble, and thus, at least under test condi-
tions, accomplishes the matching of the uncertainty of the receiver to the uncertainty of the transmitter. In this test, each of the messages is a route on the map illustrated in Fig. 4. The starting point is the position Y, heavily marked in the center of the map. Progress along the route is only to connected neighboring towns, and the route does not double back upon itself. Thus, there are four possible moves from Y - to O, J, A, or C. If the move is to O, three possibilities follow, to E, R, or H. Each message consists of six towns, as for example, that marked, O E T R V E.

The receiver may act either upon fixed-time observation or sequential observation. If he observes sequentially, the name of the town is repeated until the receiver transmits over the feedback channel a single digit denoting that he has received the signal with a satisfactory degree of confidence. By having the map available as a memory aid, the receiver can readily identify the received message as one of 972 possible routes. It is possible, however, that the received map may not be one of these routes. If it is not, it is still possible for the observer to choose the most likely of those messages actually in the ensemble. Thus, in the typical route used above, he would probably select that one as the correct message, even though he actually received O D T R V E, a route which is not in the ensemble.

Results of this test suggest that memory aids of this type can complement human memory, and indicate that such aids constitute an important area justifying extensive formal study. Such use of redundancy, of giving the receiver knowledge of the message ensemble and thus matching his uncertainty to that of the transmitter, can increase the efficiency of communications system. Further research in this area, is therefore,
FIG. 4 MAP USED AS VISUAL AID TO REDUNDANCY
clearly indicated.

2.4 Summary

Analysis of the communication's problem reveals a number of areas where further research is essential if we are to understand fully the nature of the system and the relation of the various parts to the efficient functioning of the whole. The role of the human being, whether he is on the transmitting or receiving end, is particularly significant here. As we have observed throughout, the human operator adds flexibility to a system. His presence may allow a number of parameters to remain unspecified until the system begins to operate in a particular environment. Once in the environment, the human being may adjust to the specific conditions he finds and perform more efficiently than present mechanical devices. Since he can serve so important a function, it becomes imperative to determine the effect of his behavior on the system as a whole.

From the standpoint of coding, we would like to know the effect of reading rate, volume, tone, inflections, and so forth on the efficiency of the system with a view toward determining general rules for establishing the best possible procedures for situations in which the operator may find himself. At the other end of the system, we need more information on the capacity of the human component for storing or decoding information, and on his ability to make decisions to optimize either information or utilities. We need to know also much more concerning the human being's ability to observe sequentially and to make use of redundancy in a communication's system. We would like to know the extent to which he can incorporate a priori information. Increased knowledge in all of these areas will improve our understanding of the communication's system. Such understanding will then help us in evaluating the effectiveness of any
countermeasures program.

3. THE RADAR PROBLEM

Although at first glance, the problem of the human component in a radar system would seem to be different from that in a communication system, it is in many ways much the same. To be sure, there are significant differences between radar and communication systems, differences which require that they be treated separately. Nonetheless, as the present analysis will show, insofar as the human role is considered, the similarities are greater than the differences. This analysis is based upon a specific problem: the possible influence of human performance on the range of a radar. Despite the specific nature of the problem, however, it reveals certain pertinent questions regarding human performance in a radar system in general - questions which are the same as those encountered in the discussion of the communications system.

Before the analysis can be begun, however, certain background material must be clearly understood. Since a radar system is a sensory system, the function of such a system is particularly pertinent here. Consider the block diagram shown in Fig. 5. We are concerned with the area within the dotted lines on the chart. The sensory system is what gives you the observation; the $I(x)$ (likelihood ratio) is what you get from the observation. The likelihood ratio is the ratio of two conditional probabilities: the probability of the input $x$, given signal plus noise, to the probability of the input $x$, given noise alone. It has been shown to contain all of the relevant information contained in the input $x$. The $I(x)$ computer calculates likelihood ratios and can compute only a finite number of them.
FIG. 5 BLOCK DIAGRAM OF IDEAL OBSERVER
As can be seen from this chart, the sensory system is really a part of a larger one, and its efficient operation must depend upon instructions from the larger system. That is, the larger system must determine the likelihood ratios relevant to the particular situation and time. These instructions are sent into the sensory system by the distribution computer.

In similar fashion, before we can consider the question of decisions, we must know now the sensory system fits into the larger unit. For example, if the function of the sensory system is to transmit data to a data processing center, it should be concerned with transmitting as much information as possible. It should make the decision not upon utilities, but for preservation of information. Its function should be to preserve information, not to act upon utilities. If, on the other hand, the output is used to make a decision to take action, the decision computer must consider the utilities of the decision. This problem, it will be observed, is similar to that discussed in Section 2.2 above. Finally, if the equipment is to be used for making decisions, the information fed into it from other parts of the system should be chosen by criteria for information preservation. That is, we should need the decision computer information, not decisions.

The importance of this background becomes immediately apparent when we apply it to the particular problem of the radar system. The more information fed into a sensory system from the outside, the less that has to be processed by the system itself. This fact is of the utmost importance in a radar system. A signal in noise has a capacity for carrying information, or, in other words, reducing uncertainty. Since this is true, the detection range of a radar is a function not only of the energy of the pulse return and the noise level at the receiver input, but also of the amount of information which must be processed. Thus, there is a relation between the amount of information to be processed and the range of the set.
If the \textit{a priori} uncertainty is large, the energy of the return must be correspondingly larger to reduce the uncertainty to a designated value. If, however, \textit{a priori} information has reduced the uncertainty, the energy required is less by an amount dependent upon the degree of the uncertainty. Consider, for example, Fig. 6. Here it can be seen that we can achieve a degree of certainty of, say, .98 with less energy as the number of alternatives among which the choice is made is decreased. It is clearly important, therefore, to reduce the uncertainty as much as possible before we ask the set to operate.

Suppose, for example, that our knowledge of the enemy is sufficient to ascertain that we can expect perhaps four possible targets, some 260 locations, and an infinite number of times during which the target may appear. Clearly, this is a vast amount of information, of uncertainty, to be processed. If, however, we can feed information into the set or system which will assign a high \textit{a priori} probability to one particular target, a smaller set of locations, and a limited number of times, we decrease the uncertainty, require less energy to reach a designated level of certainty, and, in effect, increase the range of the radar. In other words, we would be able to pick up the target while it is still farther out.

There are a number of ways, of course, in which this can be accomplished. A better definition of the signals that appear will certainly help. Also, when the radar is integrated into a system in which information from other sets is available, such information can be incorporated as \textit{a priori} information for the radar operator. Thus, the
FIG. 6 PROBABILITY OF A CORRECT CHOICE AS A FUNCTION OF $\sqrt{\frac{2E}{N_0}}$ WITH THE NUMBER OF ALTERNATIVES AS A PARAMETER
sensory system (the radar) must be seen in relation to a larger system; the efficiency of the sensory system will depend upon instructions from that system. Only then can we know what the system can be expected to do.

Detection, moreover, should be based, as far as is possible, on sequential observation. That is, detection should come as a result of integrated information rather than as a result of a series of individual decisions. The presence or absence of a target should not be decided at each scan of the antenna for each resolvable unit of space as displayed on the radar scope. Rather, energy should be integrated from scan to scan until it is possible to make a decision in accordance with a certain level of confidence. It is important, therefore, that information be preserved until such time as a decision must be made. Preservation of information is the most important criterion until a decision for action must be taken.

To a large extent, a consideration of all these facts affects estimation of radar range only in the way in which the performance of the decision-making device is incorporated. It is independent of the physical or electrical properties of the radar except insofar as these may be modified in the light of information being collected. The flexibility introduced with the human component is important here. Under certain conditions, the operator may elect to employ a restricted scan or restricted range, operations which can influence the total energy of returns from targets. Thus, in a radar net, another set may give information concerning a certain target, including place, direction, and speed. Such information may lead the operator to elect a particular scan or range. He will require less energy to pick up the target, and hence increase the range.
This section of the paper, therefore, will consider the important question of the effect of operator performance on radar range - an effect which will be considered in terms of fixed scan conditions, and fixed conditions for range of search. In terms of these assumptions, we will consider first the evaluation of the signal and of receiver efficiency. Such analysis will illustrate the relevant psychophysical areas for the study of radar operator performance.

3.1 The Assumed Signal Evaluation

Since operator efficiency must be seen in terms of both the efficiency of the signal and that of the receiver, we must first consider these two areas of study. The block diagram in Fig. 7 will perhaps help clarify the problem. Here, the sensory system described in the introduction to this section is incorporated into a radar system, and the three points of measurement for signal, receiver, and observer efficiency are illustrated. In other words, we need first a measure for determining signal efficiency. This factor must be modified by receiver efficiency, and the combined factor further modified by operator efficiency. In this way we can learn what performance may reasonably be expected of a radar system, for the η (efficiency) is a variable that can be dealt with practically, and is measurable in some experimental situations. There is no reason why it cannot be used in the future in increasingly complex ones.

Let us turn, then, to the question of signal evaluation to determine the basic factor in terms of which the receiver and operator efficiency must be seen. According to the theory of signal detectability, the most important parameter for such evaluation is the ratio \(2E/N_0\), where \(E\) is the signal energy and \(N_0\) is the noise power per unit bandwidth (see
FIG. 7 BLOCK DIAGRAM ILLUSTRATING POINTS OF MEASURE IN RADAR EVALUATION
Appendix). This ratio can be used to furnish the basis not only for decision tasks, but also for recognition ones. That is, when two possible signals are presented and we subtract one from the other, the difference signal we get by this means has energy which goes into the $2E/N_o$ ratio, which tells how well we can do in discriminating between the two signals. Since this is true, it is therefore assumed that the signal evaluation at least furnishes the basis for calculating an expected value of the ratio as a function of the range of the radar from the target. Once the value of $2E/N_o$ has been determined, we can work out the efficiency of the human component.

For the purpose of this discussion, a number of assumptions have been made. Figure 8 assumes a value of $2E/N_o = 1.00$ when the range is 1.00. It further assumes that the signal power of the echo varies directly with the fourth power of the reciprocal of the range, that the number of pulses returned is a constant independent of range, and that the noise at the receiver input is independent of range. No effort is made to justify these assumptions; they have been made only to furnish a basis for the demonstration which follows. The necessary evaluative procedure is not dependent upon them. One could as well assume that the noise power is increasing with range, as might be the case if countermeasures were employed.

The curves shown in Fig. 8 have taken this particular form because of the nature of the experiment upon which they are based. The situation involves two fast moving objects coming together. The lower of the two curves shows the value of $2E/N_o$ expected for each target interception. The second is the integrated value assuming that the target is intercepted once every time the range is reduced by one tenth
FIG. 8 \( \frac{2E}{N_0} \) AS A FUNCTION OF RANGE WHEN \( \frac{2E}{N_0} = 1.00 \) AT RANGE 1.00
of its original value. Because of the particular nature of the experiment, the last value every time represents the major portion of the energy of the integration, the sum of the energy to that point. If the target is intercepted more frequently, then the integrated value is increased proportionately to the frequency of interception.

Thus, if a priori information can be used to reduce the sector scanned, increasing the frequency of the intercept, the integrated value of $2E/N_0$ is expected to increase. With a more limited sector, we can sweep faster and get more interceptions; at the same time we pick up energy as well. The use of a priori information, therefore, limits the amount of information to be processed, increases the energy from the target return, and thus affects the range of the radar. Hence, from the point of view of the human operator, it is important to know the extent to which he can incorporate a priori information. Since the integrated values, moreover, imply sequential observation, the questions of the capacity of human memory and of the use of memory aids become important. The signal evaluation in terms of $2E/N_0$, therefore, raises significant psychophysical questions for further study.

These questions are pertinent, even though the assumptions underlying the particular values of $2E/N_0$ in Fig. 8 are completely arbitrary. The necessary evaluative statement remains unchanged. The curve could be established on any set of realistic assumptions, leading, of course, to a more realistic form of the curve. With realistic rather than assumed values of $2E/N_0$, we should be able to compute more accurately the efficiency of the operator.
3.2 Receiver Efficiency

Given the value of $2E/N_0$, derived from physical data, we must next modify it by a factor of receiver efficiency. This occurs at the second point of measurement shown in the block diagram in Fig. 6. To a large extent the functioning of this receiver and its efficiency will depend upon the way it is incorporated into the system and the function it is designed to serve. If there is an operator, as there is in the block diagram, the purpose of the receiver is only to transmit information, and its efficiency will depend upon the amount of information it can transmit in terms of its capacity.

If, on the other hand, there is no operator, the receiver must perform its functions, and it should act like the sensory system incorporated in the block diagram in the position of the operator. This, however, is an ideal observer; in practice, the efficiency would not be 1. Nonetheless, depending upon the function that the receiver must serve—to preserve information or to act upon utilities—we can assign at an efficiency rating ($\eta_R$) which will modify the value of $2E/N_0$ derived from the evaluation of the equipment. Once again, $\eta$ is a real value which can be used practically in the evaluation of the system.

3.3 Operator Efficiency

Once we have arrived at the value of $2E/N_0$ modified by $\eta_R$, we must next modify this combined value further by a consideration of an operator factor, $\eta_H$. This occurs at the third point of measurement in the block diagram in Fig. 6. For ease of calculation, it would be convenient to assume, as has sometimes been done, that this is a constant factor. In the past, for example, the human factor has been
assigned an efficiency of .5. Experiments indicate, however, that this factor is not a constant, and to assume that it is serves only to introduce unnecessary error.

In order to understand the variable factor \( \eta_H \) designating the efficiency of the human operator, we must first understand some specific terms \( \eta, d', \) and the relationship between these factors and \( 2E/N_0 \), since observer efficiency is related to these concepts. The definitions are based upon the block diagram shown in Fig. 9, reproduced from Tanner and Birdsall. There are four channels whose nature is determined by the position of the switches \( S_1 \) and \( S_2 \).

\( C_{11} \) is an ideal channel in which the transmitted signal is known exactly (SKE) and the receiver is ideal, that is, in relation to the particular transmitter. \( C_{12} \) is a channel in which the transmitted signal is known exactly and the receiver is one under study. \( C_{21} \) is one in which the signal is known statistically (SKS) and the receiver is ideal. (Note, however, that this is a different receiver from that in \( C_{11} \), for, as the transmitter changes, the receiver must change, since it is ideal in relation to a specific transmitter). Finally, \( C_{22} \) is a channel with the signal known statistically and the receiver one under study.

Let us first consider an experiment with Channel \( C_{12} \). \( S_1 \) is in position 1, \( S_2 \) in position 2. Signal energy \( E_{12} \) is employed, and noise power per unit bandwidth \( N_0 \) is added. The receiver's problem is to observe specified waveforms and to determine whether the waveform contains a signal plus noise, or noise alone. Performance over a large number of times can be measured in terms of a detection rate \( P_{SN}(A) \)—the probability that if a signal was presented, it was
accepted,—and a false alarm rate $P_N(A)$—the probability of noise alone being accepted. The "hit rate" is therefore $P_{SN}(A)/P_N(A)$.

The second experiment is a mathematical calculation. It is similar to the first except that $C_{11}$—that is, an ideal receiver, is employed. With $S_2$ in position 1, therefore, the experiment is repeated, with the energy of the signal attenuated until the performance attained in the previous experiment is matched. This energy is $E_{11}$. The efficiency of the receiver under study for SKE ($\eta_R$), therefore, may be defined as $E_{11}/E_{21}$, where $E_{11}$ is the energy level at which the performance is matched. The measure $d'$ is then defined by the equation

$$(d')^2 = \eta_R \frac{2E_{12}}{N_0} = \frac{2E_{11}}{N_0}$$

That is, $(d')^2$ is that value of $2E/N_0$ necessary to lead to the observed performance, given SKE and its ideal receiver.

Let us next consider a second pair of experiments, both of which are mathematical calculations. The channel employed is $C_{21}$; that is, the signal is known statistically and the receiver is an ideal one for that statistical ensemble. Energy $E_{21}$ and noise $N_0$ are employed, and a performance measure is established. $S_1$ is then moved to position 1, changing the channel to $C_{11}$ (and, it should be noted, also changing the ideal receiver). Energy is attenuated until performance is matched at $E_{11}$. In the case of the signal known statistically, this permits the calculation of the efficiency of the transmitter, $\eta_t$, for the signal known statistically (SKS) as $E_{11}/E_{21}$. At this point, it should be obvious that if we proceed in this way with paired experiments, we can establish a receiver efficiency for any SKS, which might be expressed as $\eta_R(SKS) = E_{21}/E_{22}$.
One final definition remains to be made—\( \eta \), that is, operator efficiency. Since \((d')^2\) is that value of \(2E/N_o\) needed if we have ideal conditions, for a particular experiment \( \eta = ((d')^2)/(2E/N_o) \), where \( E \) is the energy used in the experiment. The relationship, therefore, is one between the energy used under ideal conditions and that required in a real situation.

With these definitions behind us, we can turn to the actual calculation of \( \eta_H \), the efficiency of the human operator as it affects the performance of the system. In visual experiments, Tanner and Swets observed that \( d' \) for weak signals varied approximately as the square of signal intensity (energy). If the signal were known exactly, however, \( d' \) should vary as the square root of signal intensity. Since \((d')^2 = 2E/N_o\), this says that observer efficiency varies as a function of \((2E/N_o)^{2}\) for weak signals.

Although this information is important in the present context, the data acquired from the visual experiments are not in usable form, for in these experiments, the values of \(2E/N_o\) were not known precisely. In order to put these data into usable form for the purposes of this paper, we must extrapolate from the results of auditory experiments. In these experiments, the values of \(2E/N_o\) are known precisely and we can calculate the \( \eta \)'s. Such extrapolation is, of course, very dangerous. It is done here only with the purpose of establishing an approximation. Quite obviously, further visual experiments and more precise information are necessary before we can be sure of the accuracy of the values of \( \eta \).

The extrapolation is based upon the following theoretical observations: (1) signals known statistically, as they get large, approach in \((d')^2\) signals known exactly, and (2) observer efficiencies in auditory experiments involving large signals approach values as high
as .5 or greater. On the basis of these observations, it is assumed that when $2E/N_o = 10$, the observer efficiency is .5 and becomes lower as $2E/N_o$ decreases, in accordance with the observation reported by Tanner and Swets. This is, it must be repeated, an extrapolation which should be checked experimentally. It may well be too optimistic.

These estimates are presented in Table I. The $\eta$ is derived as a result of the ratio $(d')^2/(2E/N_o)$ as defined above. The data in these columns are shown graphically in Fig.10, where the estimate of the human observer's performance is compared to the ideal. Once these values of $\eta$ are accurately determined, we can proceed to modify the combined evaluation of the signal and receiver efficiency. The value of $2E/N_o$ as derived from signal evaluation and as modified by the efficiency of the receiver ($\eta_R$) would then be further modified by the operator efficiency factor ($\eta_H$) for that $2E/N_o$ as estimated in Table I.

3.4 The Capacity of a Signal to Lead to a Correct Choice Among One of M Alternatives

Now that the general process is understood by which the capacity of the system may be modified by factors of receiver and operator efficiency, we may return to the question of the evaluation of the equipment as presented in Section 3.1 above, to show specifically how the efficiency may be affected by the use of a priori information.

In a short unpublished memo, Peterson and Birdsall (Ref. 6) have considered the problem of making a correct choice from M equally-likely orthogonal signals. If we consult the two statistical distributions basic to the calculations of $2E/N_o$ (see Appendix), this is a case in which M-1 observations are from the distribution conditional upon noise
<table>
<thead>
<tr>
<th>$\frac{2E}{N_0}$</th>
<th>$(d')^2$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
<td>.5000</td>
</tr>
<tr>
<td>9</td>
<td>4.1</td>
<td>.455</td>
</tr>
<tr>
<td>8</td>
<td>3.4</td>
<td>.425</td>
</tr>
<tr>
<td>7</td>
<td>2.8</td>
<td>.400</td>
</tr>
<tr>
<td>6</td>
<td>2.0</td>
<td>.333</td>
</tr>
<tr>
<td>5</td>
<td>1.4</td>
<td>.280</td>
</tr>
<tr>
<td>4</td>
<td>.9</td>
<td>.225</td>
</tr>
<tr>
<td>3</td>
<td>.53</td>
<td>.177</td>
</tr>
</tbody>
</table>

Table I. Estimation of Operator Efficiency

<table>
<thead>
<tr>
<th>M</th>
<th>0</th>
<th>.5</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>3.0</th>
<th>4.0</th>
<th>5.0</th>
<th>6.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>.50</td>
<td>.64</td>
<td>.76</td>
<td>.92</td>
<td>.984</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.33</td>
<td>.48</td>
<td>.63</td>
<td>.87</td>
<td>.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.25</td>
<td>.39</td>
<td>.55</td>
<td>.83</td>
<td>.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.125</td>
<td>.24</td>
<td>.39</td>
<td>.71</td>
<td>.92</td>
<td>.987</td>
<td>.999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.0625</td>
<td>.14</td>
<td>.26</td>
<td>.42</td>
<td>.60</td>
<td>.87</td>
<td>.975</td>
<td>.997</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>.03125</td>
<td>.08</td>
<td>.17</td>
<td>.31</td>
<td>.48</td>
<td>.80</td>
<td>.957</td>
<td>.995</td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>.00391</td>
<td>.02</td>
<td>.05</td>
<td>.22</td>
<td>.57</td>
<td>.866</td>
<td>.979</td>
<td>.9984</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>.0010</td>
<td>.01</td>
<td>.02</td>
<td>.12</td>
<td>.41</td>
<td>.765</td>
<td>.951</td>
<td>.9950</td>
<td></td>
</tr>
</tbody>
</table>

Table II. The Probability of a Correct Choice of One of M Equally Likely Orthogonal Alternatives (from Peterson and Birdsall).
FIG.10 ESTIMATE OF HUMAN OBSERVER'S PERFORMANCE COMPARED TO THE IDEAL
alone, and one observation is from the distribution conditional upon signal plus noise. The probability of a correct choice is the probability that the observation from signal plus noise is greater than the greatest of the M-1 observations from noise alone. For the ideal case, this is given in the following equation:

$$P(c) = \int_{F(x)}^{x} f(x - \frac{2E}{N_0}) \, dx,$$

where $F(x) = \int_{0}^{x} f(x) \, dx$, and $f(x)$ is the probability density for the observation $x$.

Peterson and Birdsall have constructed a table based on an approximation of this equation, reproduced here as Table II, and shown graphically in Fig. 6. From their approximations, Fig. 11 has been constructed. This figure shows the value of $2E/N_0$ necessary to lead to a correct decision as a function of the a priori uncertainty ($\log_2 M$) with performance criteria of .85, .90, and .95 probabilities. As can be seen from the graph, the larger the a priori uncertainty, the greater the energy needed to reduce that uncertainty to a designated value.

3.5 Estimating M

It is assumed here that if the error of estimate of target location is normally distributed over a single dimension, then the uncertainty ($\log_2 M$) of target location is approximately the same as in the case of a target equally likely over the equivalent rectangular distribution. An approximation of $M$ is then given by

$$M = \frac{\sigma x}{2 \text{ Antenna Beam Width}}$$

where $\sigma$ is the estimate of error in a priori estimate of target position.
FIG. 11 $\frac{2E}{N_0}$ REQUIRED TO MEET CRITERION AS A FUNCTION OF UNCERTAINTY
3.6 Summary

Although the radar problem is in many ways fundamentally different from the communications problem, the analysis of the radar range problem presented above illustrates certain fundamental similarities as far as the human component is concerned. The relevant psychophysical areas for the study of radar operator performance are not fundamentally different from those illustrated in the communications problem. In both, the relevant questions are the same. Just as the use of redundancy in the communications problem raised the question of the ability of the human being to incorporate a priori information, so also does the radar problem, insofar as such information affects the energy requirements, and hence the range, of the system.

Thus, we may summarize the areas where further study is required by means of the following questions: (1) Can the operator incorporate a priori information? (2) What is the extent of his memory, and what kind of aids can be introduced to supplement this memory? (3) Can the operator act as a sequential observer? (4) Can the operator optimize information? These questions are very similar to those asked in the communication problem. All will require further study before we can be sure of just how the human component may be incorporated into a system and how his incorporation will affect the performance of the system as a whole.
4. THE COUNTERMEASURES PROBLEM

Once the fundamental natures of the communications and radar problems are fully understood and the pertinent areas requiring further study have been isolated, we can turn finally to the countermeasures problem itself. All of the questions we have treated in the preceding two chapters are pertinent here. Most especially, the dominant idea which has underlain much of the foregoing discussion needs to be stressed again: that the efficiency of the system depends to a great extent on the specific job the system is designed to perform so that the more clearly we can specify the situation in which the system will act, the more accurately we will be able to evaluate its efficiency. The role of each of the factors in the system must be carefully studied for two important reasons. First, we must clearly understand the effect each factor—including the human one—has on the system as a whole so that we can accurately evaluate the system. Second, once given this information, we must strive toward the formulation of a more general rule in terms of which the system may be evaluated.

The importance of this second concept cannot be overemphasized. Although we must study the system in terms of specific situations, the data we acquire may not be entirely useful in the evaluation of equipment if that equipment should be used for an entirely different purpose. Since the efficiency depends to a great extent on the particular situation or game, a change in that situation may render the data derived from the first situation all but useless. Clearly,
what we must seek is a more general rule, a more general means of evaluation which can be used to cover a variety of specific situations. The data should be in some form—still to be determined—which will not be designed to tell the user the specific evaluation of the equipment, but which will enable him to calculate that efficiency for himself given the specific situation in which he is engaged. It is this second concept which is our major concern in this chapter.

As in the preceding two sections, the analysis of the countermeasures problem must proceed from a realization that the problem is derived from the fact that we are involved with a system. As such, it can best be understood in terms of the material that we have already covered. Consider the relation first to the communications problem. If we consult the block diagram in Fig. 12, illustrating the basic countermeasures problem, we can see immediately that what we have here is really a modified version of Shannon's fundamental diagram of a communication channel (Fig. 1), with the operator and jammer added. As we would expect, therefore, the fundamental questions of information, channel capacity, entropy, and so forth, all apply here, especially insofar as they are affected by the presence of the jammer. In a similar fashion, the psychophysical questions raised by the incorporation of a human component are equally pertinent in this context.

Certain similarities will also be noted between this diagram and the one basic to the radar problem shown in Fig. 6, wherein the sensory system was incorporated into the radar. Of the greatest importance are the three points of measurement used to determine the effectiveness of jamming. These correspond generally with the three points of measurement for determining equipment, receiver, and operator
efficiency in the earlier problem and may be easily understood in
terms of that problem. Thus, the question of determining the value
of $\eta$, basic to the radar problem, is pertinent to the countermeasures
problem as well, and has to be worked out for each factor in the
system, including the human one. The diagram in Fig. 12, therefore,
is presented here to assist the reader, familiar with the preceding
problems, in identifying the factors involved in evaluating the
effectiveness of a jamming tactic.

The factors involved in such evaluation may be divided into
two main categories, two of which relate to the communications
system in general, and a third which is related to the countermeasures
problem in particular. The first and second categories may be dis-
tinguished from each other according to their relation to the specific
situation. The word "specific," however, is used here in a somewhat
unusual sense. It does not refer to the particular physical parameters
involved—such as distance or terrain—but rather to those elements in
the situation determined by the particular game being played. The
first group of factors is specific to the game; the second is not.

Those factors, among others, which are specific to the
particular situation may be listed as follows:

(1) Capacity required of the forward channel.
(2) Coding scheme employed.
(3) Decoding scheme employed, including such factors as
    the effectiveness of the operator.
(4) Use of the feedback channel.
(5) Criterion of acceptable performance, such as
    (a) Permissible error rate,
(b) Permissible miss rate,
(c) Importance of time.

All of these areas are familiar to the reader from the discussion of the communications problem in Section I above, and a number of the problems concerning at least some of them have been touched upon there. They should require, therefore, no additional discussion here.

The physical factors involved—those not considered specific to the game—are as follows:

(1) Capacity of the forward channel.
(2) Distances involved.
(3) Atmospheric attenuation.
(4) Terrain factors.
(5) Condition of equipment.
(6) Regulation of power supply.
(7) Susceptibility of equipment, such as
   (a) Saturation,
   (b) Effect of non-linear elements,
   (c) Detectability by jammer.

The first six of these factors are also important in the third category, that which, as we have noted, relates most particularly to the countermeasures problem.

This third category of factors, somewhat different from the previous two in that it relates to the jamming system rather than to the general communications system, includes the first six physical factors listed above, and adds two others:
(1) The ability to take advantage of the detectability of the jamming target, and

(2) The ability to transmit an adequate jamming signal.

Once these three main sets of factors are understood, we may proceed to the question of the evaluation of a jamming operation.

As in the case of the radar problem, the effectiveness of the system cannot be measured at only one point, since to measure the effectiveness in this way fails to take into account the relative importance of each of those factors which contribute to the evaluation, and leaves the observer in the dark concerning the relation among those factors. For this reason, in the block diagram in Fig. 12, three points of measurement are indicated—for signal jamming, receiver jamming, and observer jamming—points which do permit the isolation of the relevant factors and which should furnish a much sounder foundation for collecting information upon which to base future research and development programs. Let us examine each of these points in turn, therefore, to see the relation among them and the relative effect each has on the valuation of the countermeasures system as a whole.

The first point of measurement comes before the signal enters the receiver and is the point for measuring signal jamming. As defined by Hok, this is actually a measure of the ability to reduce the capacity of the channel up to the input of the receiver. Remember that capacity is defined in terms of the information (in the special sense of communication theory) which the channel can transmit. The jamming signal reduces that capacity and increases the uncertainty with which the signal is received. The measure is made at this point rather than at the second point because any but an ideal receiver
will further increase that uncertainty. An ideal receiver uses all the information at the input, and were such a one employed, then the measure at point one (for signal jamming) would be the same as that at point two (for receiver jamming). In an actual case, however, the receiver will add additional uncertainty, and we wish to isolate the first important factor. This factor is a measure of the additional entropy (the degree of randomness) of the input signal as a result of the jamming signal.

The second point of measurement is that for receiver jamming. What we wish to isolate here is the degree of uncertainty which comes as a result of the receiver, that is, of qualities inherent in the equipment itself. Measurement at this second point requires, therefore, a measure of the additional entropy at this point, and then the isolation of that part of the additional entropy which is due specifically to the receiver. Because realizable receivers have a finite range and non-linear elements, the efficiency of the receiver may depend upon conditions at the input. Consequently, if the efficiency of the receiver changes as the signal is jammed, then the receiver is contributing to the entropy beyond that of signal jamming. Only with an ideal receiver are the measures at points one and two the same. To measure only at point two, therefore, will give a false estimate of the efficiency of signal jamming. Similarly, a measure at only point one would ignore the fact that realizable receivers add to the uncertainty with which the signal is received. Measures at both points, therefore, are required if we are to get an accurate estimate of the efficiency of the jamming tactic.
This concept will perhaps be further clarified if we consider it in terms of a specific example. Consider a channel, the capacity of which is measured in terms of bandwidth and the signal-noise ratio, where only noise relevant to the signal is considered in calculating the ratio. Designate this capacity as $D_o$. Now if the measure is made at the input of the receiver, that is, at our first point of measurement, the effectiveness of signal jamming can be expressed by the following ratio:

$$\frac{D_o - D_j}{D_o}$$

where $D_o$ is the capacity of the channel without jamming and $D_j$ is the capacity with jamming. We can learn thereby the degree by which the jamming signal has reduced the capacity of the channel. This gives us, then, a measure of the effectiveness of the signal jamming.

Let us proceed to the second point of measurement, that for receiver jamming. If the measure is made here, we can express the effectiveness by the following ratio:

$$\frac{\eta_o D_o - \eta_j D_j}{\eta_o D_o}$$

where $\eta_o$ is the efficiency of the receiver when the input signal is $D_o$ and $\eta_j$ is its efficiency when the input signal is $D_j$. What we have done in effect is to modify the measure of the effectiveness of the signal jamming by a factor of receiver efficiency to arrive at an overall estimate of the efficiency of the jamming operation to this point.

This last expression can be rewritten as follows:

$$(1 - \frac{D_j}{D_o}) + \frac{D_j}{D_o} (1 - \frac{\eta_j}{\eta_o})$$

where \(1 - \frac{\eta_j}{\eta_o}\) can be considered the effectiveness of the receiver.
jamming. This rewriting of the formula is useful because it isolates the term \(1 - \frac{\eta_j}{\eta_0}\). If one is interested in the susceptibility of the receiver, this is the factor which should be studied, because it expresses that part of the effectiveness of the jamming operation that comes as a result of qualities inherent in the receiving equipment itself.

Let us proceed further to the third point of measurement, that for system jamming. The third factor which enters here is the efficiency of the human operator, a factor which, as we have seen in the radar problem in Chapter 3, modifies further the efficiency of the system as a whole. To account for this third factor, let us extend the measure as it is rewritten in the paragraph just above. It can be extended to n terms. If we extend it to three terms to take in the factor of operator efficiency, the measure reads:

\[
(1 - \frac{D_j}{D_0}) + \frac{D_j}{D_0} \left(1 - \frac{\eta_{RJ}}{\eta_{RO}}\right) + \frac{D_j\eta_{RJ}}{D_0\eta_{RO}} \left(1 - \frac{\eta_{HJ}}{\eta_{HO}}\right)
\]

where the \(\eta\)'s are distinguished by their subscripts as \(\eta_R\), the efficiency of the receiver, and \(\eta_H\) the efficiency of the human operator. Just as the second step in the measurement process isolated a term useful in studying the susceptibility of the equipment, this expression isolates the term \(1 - \frac{\eta_{\text{HS}}}{\eta_{\text{HO}}}\), which can be considered the effectiveness of observer jamming. This would include the additional uncertainty, over and above that caused by signal and receiver jamming, which can be attributed to the human operator. This then is the factor which should be studied if one is interested in the susceptibility of the human being. It expresses that part of the jamming operation that comes about because of the nature of the human operator.
The total expression, which takes into account all three factors, describes the effectiveness of the system jamming. Each of the factors in parenthesis represents the effectiveness at one of the points of measurement: the signal jamming, the receiver jamming, and the observer jamming. Any measure which considers only the system as a whole fails to isolate these three factors which contribute to the specific effects in any single test. On the other hand, measures taken at each of three points specified in the block diagram in Fig. 10 do permit their isolation. They provide a basis for studying individual factors, such as equipment evaluation and observer efficiency. They should furnish, therefore, a more solid foundation for the collection of information upon which future research and development programs may be based.

The form in which these data should be presented, however, is as we have observed, still to be determined. It seems obvious that in each stage of the measurement, the results of the measures are, to some extent at least, dependent upon the particular situation or game. The $\eta$'s are certainly variable factors, their values varying with the specific situation. This specificity must be realized and studied, not only so that we can understand the role of each factor in a specific game, but more especially so that we can eventually determine more general means of evaluation, means which can be used for a variety of specific situations.

A concrete example will perhaps clarify the issue here. What we are seeking are data which can be presented in a way similar to that in which a manufacturer presents specifications on an oscilloscope. He does not present a number which evaluates the equipment in terms of
a particular use the customer has for the product. Rather, he presents a set of data which permits the customer to calculate the value of the product for his own use. The data are designed to permit this calculation for a large number of uses. In like manner, we must find a means for presenting data in a way that will permit the evaluation of not just one jamming operation. Rather, the data should be in such form as will permit those engaged in countermeasures programs to evaluate their own specific situations for themselves.

We should seek, therefore, a general rather than a specific rule. Once we are aware of this fact, we can state the fundamental problem. Laboratory data on jamming tests should be presented in this way: there should be a set of specifications permitting their use in evaluating a large number of specific situations, rather than a number evaluating a specific situation. What the specifications should be, and the ways in which one uses them to estimate the results in a specific situation, are closely related questions. The answers to them depend, however, on knowledge in areas which are still relatively undeveloped. As has been stressed throughout this paper, the function of the human component in any system is one such area.

Particularly significant is an understanding of the way in which the human being introduces flexibility into the total system, since his ability to incorporate a priori information, to function as a sequential observer, or to optimize information or utilities can have a significant effect upon the functioning of any system. Furthermore, his self-evaluating and self-adjusting qualities may lead to such flexibility that he might even be able to offset, at least partially, the effect of jamming in a communications system. The
relation of the human component to the system as a whole and his possible effect on its efficiency are, therefore, areas where considerable research is certainly warranted.
APPENDIX

MATHEMATICAL CALCULATION OF THE DISCRIMINABILITY OF TWO SIGNALS

The purpose of this appendix is to describe the mathematical development of the determination of the restrictions placed on performance in a detection experiment by the environmental conditions.

A signal is defined here as a voltage waveform. In a well-defined time interval, one of two signals perturbed by noise is presented to an observer or a receiver. It is the observer's task to state which of the two signals was contained in the input waveform.

In this analysis, the voltage waveforms, the signals, are precisely defined functions of time, $S_1(t)$ and $S_2(t)$, entirely contained within the observation interval, $t$ to $t + T$. The precise definition of the waveform means that, if the waveform exists in the interval, the voltage amplitude at any instant in time, $t_i$, is a precisely specified value, $S_{1i}$.

The perturbing noise is assumed to be Fourier series band-limited white Gaussian noise which is added to the signal. This set of assumptions is made to permit discrete statistical analysis. Since the noise is series band-limited, a noise waveform within the interval can be precisely specified by $2WT$ measures of voltage amplitude where $W$ is the bandwidth and $T$ is the duration of the interval. The fact that the noise is white implies that the power density is uniform at every frequency within the band $W$. The Gaussian assumption requires the voltage

---

1 The mathematical development presented here follows closely that of Peterson, Birdsall and Fox (Ref. 7).
amplitude of the noise to be a normal variate with mean zero and variance, 
or noise power, N. The assumption that the noise is added to the signal
states that at each instant of time in the interval, the voltage of the
input waveform, x(t), is the sum of two voltages, the signal voltage and
the noise voltage.

The purpose of this appendix is to show how well an optimum
receiver can do in the task of specifying which of the two waveforms,
S_1(t) or S_2(t), is contained in an input waveform, x(t), when the
receiver is faced with the restrictions defined by the assumptions out-
lined above. The analysis treats the receiver as testing statistical
hypotheses. Measures of performance are statements of averages expected
of the receiver over an infinite sequence of independent observations.

It has been shown that an optimum receiver bases its decisions
on a likelihood ratio criterion (Ref. 7). That is to say, the optimum
receiver accepts one of the hypotheses whenever the likelihood ratio is
a value greater than a weighting function, \( \beta \), otherwise it accepts the
other. If monotonic transformations of both likelihood ratio and the
weighting function are incorporated, an equivalent decision rule can be
developed. The task of the analysis is to study the distribution either
of the likelihood ratio or some appropriate monotonic function of the
likelihood ratio first under the condition that one of the signals,
S_1(t), exists; and then under the condition that S_2(t) exists.

In order to perform the analysis, a sampling theorem based on
the series band-limited assumption is employed. The purpose of the
sampling theorem is to permit the use of discrete statistics. It says
essentially that the input waveform, the noise, and the signal can each
be specified completely by 2WT independent voltage amplitudes \[ x(t) = \]
\[ x_1, x_2, \ldots, x_{2^{WT}}; S_{1}(t) = S_{1,1}, S_{1,2}, \ldots, S_{1,i}, \ldots, S_{1,2^{WT}}; n(t) = n_1, n_2, \ldots, n_i, \ldots, n_{2^{WT}}. \] From this it follows that

\[
2^{WT} \sum_{i=1}^{t+T} (S_{1,i})^2 = 2 \int_{t}^{t+T} [S_{1}(t)]^2 \, dt = 2WE_1
\]

where \( E_1 \) is the energy in the voltage waveform, \( S_1(t) \), assuming that
the waveform exists over a one ohm resistance.

The likelihood ratio is defined as

\[
l[x(t)] = \frac{f_{S_1}(t)[x(t)]}{f_{S_2}(t)[x(t)]}
\]

where \( f_{S_1}(t)[x(t)] \) and \( f_{S_2}(t)[x(t)] \) are probability densities conditional upon the waveform \( x(t) \) resulting from \( S_1(t) \) and \( S_2(t) \) respectively. Since the \( 2^{WT} \) points are independent and specify the waveform in its entirety,

\[
f_{S_1}(t)[x(t)] = \prod_{i=1}^{2^{WT}} f_{S_{1i}}(x_i)
\]

and

\[
f_{S_2}(t)[x(t)] = \prod_{i=1}^{2^{WT}} f_{S_{2i}}(x_i)
\]

Since \( n_1(t) \) is a normal variate, \( x_i \) is a normal variate with mean \( S_{1i} \) or \( S_{2i} \) and variance \( N \). Therefore

\[
f_{S_{1i}}(x_i) = \frac{1}{2\pi N} \left( x_i - S_{1i} \right)^2
\]

79
\[ f_{S_{21}}(x_i) = \frac{1}{2\pi N} \cdot \frac{-(x_i - S_{21})^2}{2N} e \]

Substituting (4) in (3), and the result in (2) leads to the following equation:

\[ l[x(t)] = \frac{\frac{-\Sigma(x_i - S_{11})^2}{2N} - \frac{-\Sigma(x_i - S_{21})^2}{2N}}{e} \] (5)

It is now appropriate to consider the natural logarithm of the likelihood ratio;

\[ \ln l[x(t)] = \frac{-\Sigma x_i^2 + 2\Sigma x_i S_{11} - \Sigma S_{11}^2 + \Sigma x_i^2 - 2\Sigma x_i S_{21} + \Sigma S_{21}^2}{2N} \] (6)

\[ \ln l[x(t)] = \frac{2\Sigma x_i S_{11} - \Sigma S_{11}^2 - 2\Sigma x_i S_{21} + \Sigma S_{21}^2}{2N} \]

Examination of Eq. (6) shows that \( x_i \) is the only variable. Since at each of the \( i \) points \( x_i \) is a normal variate regardless of which signal is present, the \( \ln l[x(t)] \) being a sum of independent normal variates is likewise a normal variate.

It remains only to determine the means and variances of the two distributions conditional upon the inclusion first of \( S_{1}(t) \) in \( x(t) \) and then of \( S_{2}(t) \) in \( x(t) \).

If \( S_{1}(t) \) is included, then the expected value of each \( x_i \) is \( S_{1} \). Substituting this in Eq. (6) and letting \( M_{1} \) be the mean of the \( \ln l(x) \) conditional upon the inclusion of \( S_{1}(t) \) leads to the result:

\[ M_{1} = \frac{2\Sigma S_{11}^2 - \Sigma S_{11}^2 - 2\Sigma S_{11} S_{21} + \Sigma S_{21}^2}{2N} \]
\begin{equation}
M_1 = \frac{\sum S_{1i}^2}{2N} + \frac{\sum S_{2i}^2}{2N} - \frac{\sum S_{1i}S_{2i}}{N} \tag{7}
\end{equation}

Defining \( N_0 = \frac{N}{W} \) and employing the sampling theorem (1), Eq. (7) becomes

\begin{equation}
M_1 = \frac{E_1}{N_0} + \frac{E_2}{N_0} - \frac{\sum S_{1i}S_{2i}}{N} \tag{8}
\end{equation}

The last term is a correlation term.

\[ \rho = \frac{\sum S_{1i}S_{2i}}{\sqrt{\sum S_{1i}^2} \sqrt{\sum S_{2i}^2}} \]

\begin{equation}
\frac{\sum S_{1i}S_{2i}}{N} = \sqrt{\frac{2WE_1}{N}} \sqrt{\frac{2WE_2}{N}} \rho \tag{9}
\end{equation}

\[ = \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}} \rho \]

substituting Eq. (9) in Eq. (8)

\begin{equation}
M_1 = \frac{E_1}{N_0} + \frac{E_2}{N_0} - \rho \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}} \tag{10}
\end{equation}

By examination, it can be seen that the mean \( M_2 \) of the distribution conditional upon the inclusion of \( S_2(t) \) is the negative of \( M_1 \):

\begin{equation}
M_2 = -\frac{E_1}{N_0} - \frac{E_2}{N_0} + \rho \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}} \tag{11}
\end{equation}

The difference between the means is

\begin{equation}
d = \frac{2E_1}{N_0} + \frac{2E_2}{N_0} - 2\rho \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}} \tag{12}
\end{equation}
Now consider the variance $\sigma_1^2$ of Eq. (6) conditional upon the existence of $S_1(t)$. The terms $\sum_{s_{1i}}$ and $\sum_{s_{21}}$ are constants and consequently contribute no variance.

$$\sigma_1^2 = \frac{(\sum_{x} S_{1i} - \sum_{S_{21}})^2}{4N^2} \quad (13)$$

Since the $i$ points are independent and values of $S_{1i}$ and $S_{21}$ are constant for each $i$, this expression can be rewritten

$$\sigma_1^2 = \frac{\sum_{S_{1i}} (\Sigma S_{1i} - \Sigma S_{21})^2}{N^2}$$

The expected value of $\sum x_i^2$ is $N$ by definition, and therefore

$$\sigma_1^2 = \frac{\sum S_{1i}^2}{N} + \frac{\sum S_{21}^2}{N} - \frac{2E_{1i}S_{21}}{N} \quad (14)$$

Employing the sampling theorem

$$\sigma_1^2 = \frac{2E_1}{N_0} + \frac{2E_2}{N_0} - 2\rho \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}} \quad (15)$$

It is obvious that

$$\sigma_2^2 = \sigma_1^2 \quad (16)$$

Thus the two distributions are normal with equal variance.

The value of $d'_{\text{opt}}$, for the optimum receiver is the difference in the means divided by the standard deviation:

$$d'_{\text{opt}} = \sqrt{\frac{2E_1}{N_0} + \frac{2E_2}{N_0} - 2\rho \sqrt{\frac{2E_1}{N_0}} \sqrt{\frac{2E_2}{N_0}}} \quad (17)$$

82
It should be pointed out that \( \rho \), a correlation term, describes the correlation between the mathematical description of the two signals. In the case considered in the text, the two signals are pulses of sine waves, differing only in amplitude. In this case \( \rho = 1 \). Therefore

\[
 d'_{\text{opt}} = \sqrt{\frac{2E_1}{N_0}} - \sqrt{\frac{2E_2}{N_0}} = \sqrt{\frac{2E_\Delta}{N_0}} \tag{18}
\]

where

\[
 E_\Delta = \int_{t}^{t+T} [S_1(t) - S_2(t)]^2 \, dt.
\]

It is necessary to point out again that the result depends on the particular set of assumptions described at the beginning of the appendix. If the assumption of series band-limit had been different, the result would have been different. For example, if the noise is assumed to be transform band-limited, and therefore analytic, Slepian (Ref. 8) has shown that the signal is perfectly detectable since in this case the noise is deterministic from \( -\infty \) to \( +\infty \). Proofs of perfect detectability require mathematically precise measurement. Any error of measurement, no matter how small, leads to the proof blowing up. So far the assumptions employed indicate that if the signal is finitely detectable, the series band-limit assumption at least leads to a result nearly that of any other set of assumptions so far examined. The Fourier series band-limited assumption is therefore accepted as adequate for the present.

To support the assumptions of white and Gaussian, one need only rely upon the description of the General Radio 1390A noise generator
which have been repeatedly checked. The additivity assumption is supported by the fact that the noise and signal are combined in an additive electronic network.
REFERENCES


